

# Joint Source Channel Coding with Side Information Using Hybrid Digital Analog Codes

Makesh Pravin Wilson  
Dept. of Electrical Engineering  
Texas A&M University  
College Station, TX 77843  
makesh@ece.tamu.edu

Krishna Narayanan  
Dept. of Electrical Engineering  
Texas A&M University  
College Station, TX 77843  
krn@ece.tamu.edu

Giuseppe Caire  
Dept. of Electrical Engineering  
University of Southern California  
Los Angeles, CA 90089  
caire@usc.edu

## Abstract

We study the joint source channel coding problem of transmitting an analog source over a Gaussian channel in two cases - (i) the presence of interference known only to the transmitter and (ii) in the presence of side information known only to the receiver. We introduce hybrid digital analog forms of the Costa and Wyner-Ziv coding schemes. Our schemes are based on random coding arguments and are different from the nested lattice schemes by Kochman and Zamir that use dithered quantization. We also discuss superimposed digital and analog schemes for the above problems which show that there are infinitely many schemes for achieving the optimal distortion for these problems. This provides an extension of the schemes by Bross et al to the interference/side information case. We then discuss applications of the hybrid digital analog schemes for transmitting under a channel signal-to-noise ratio mismatch and for broadcasting a Gaussian source with bandwidth compression.

## I. INTRODUCTION AND PROBLEM STATEMENT

Consider the classical problem of transmitting  $K$  samples of a discrete-time independent identically distributed (i.i.d) real Gaussian source  $\mathbf{v}$  in  $N$  uses of an additive white Gaussian noise (AWGN) channel such that the mean-squared error distortion is minimized. Let the source be encoded into the sequence  $\mathbf{x}$  which satisfies a power constraint  $E[\mathbf{x}\mathbf{x}^T] \leq NP$ . Let us first consider the case of  $K = N$  and let the output of the AWGN channel  $\mathbf{y}$  be given by

$$\mathbf{y} = \mathbf{x} + \mathbf{w}$$

where  $\mathbf{w}$  is a noise vector of i.i.d Gaussian random variables with zero mean and variance  $\sigma^2$ . If the source variance is  $\sigma_v^2$ , then the optimal mean-squared error distortion that can be achieved is  $D_{opt} = \frac{\sigma_v^2}{1 + \frac{P}{\sigma^2}}$ . This optimal performance can be achieved by two very simple schemes. The first one is separate source and channel coding, where the source is first quantized and the quantization index is transmitted using an optimal code for the AWGN channel. The second scheme is uncoded (analog) transmission with power scaling [3, 4], where the source is not explicitly quantized. Recently, it was shown by Bross, Lapidot and Tinguely [6] that there is a family of infinitely many schemes that are optimal, which contains the separation based scheme and uncoded (analog) transmission as special cases.

This work was supported by the National Science Foundation under grant CCR 0515296.

In this paper, we consider the problem of transmitting  $K$  samples of an i.i.d Gaussian source through  $N$  uses of an AWGN channel. We will refer to the ratio of  $N/K$  as the bandwidth efficiency  $\lambda$ . We first consider the case of  $\lambda = 1$  and the presence of an interference known only to the transmitter and/or side information available only at the receiver. We derive hybrid digital analog (HDA) coding schemes for these cases where the source is not explicitly quantized and show that they can obtain the optimal distortion. These can be viewed as the equivalent of uncoded transmission but in the presence of an interference or side information. Then, we show that there is a family of infinitely many schemes that are optimal for this problem which contain pure separation based schemes and HDA schemes as special cases. This can be viewed as the extension of Bross, Lapidoth and Tinguely's [6] result in the presence of interference/side-information. An interesting aspect of the hybrid digital analog coding schemes proposed here is that they do not require binning unlike their separation based counterparts.

The HDA scheme proposed here for the case of interference known at the transmitter is closely related to the scheme considered by Kochman and Zamir in [7], although this was developed independently. The difference is that the proposed scheme is based only on random coding arguments and does not use nested lattices like in [7]. As a result, the relationship between the auxiliary random variable and the source is made more explicit. We also consider several applications of the HDA schemes which are not considered in [7]. We consider the non-asymptotic SNR case unlike in [7]. Further, the performance of this scheme in the presence of SNR mismatch is analyzed. Finally, the use of a HDA Costa based scheme for broadcasting a Gaussian source to two users with bandwidth compression, where  $\lambda < 1$  is discussed. In the case of side-information available only at the receiver, the proposed scheme is similar to the scheme in [2] and again uses random coding arguments instead of nested lattices.

The paper is organized as follows. First in Section II, we discuss the problem of transmitting an i.i.d Gaussian source in the presence of a Gaussian interference known only to the transmitter. We introduce a hybrid digital analog (HDA) Costa coding scheme where the source is not explicitly quantized and show that this is optimal. We then discuss a generalized HDA Costa coding scheme and show that there infinitely many schemes that are optimal. In Section III, we discuss similar schemes and results for the case of having side information available only at the receiver (Wyner-Ziv problem) and in Section IV, briefly consider the situation having interference known only to the transmitter and side-information available only at the receiver. In [11], Merhav and Shamai have shown that separate Wyner-Ziv coding followed by Gelfand-Pinsker coding is optimal for this problem. However, we show that there is a joint source-channel coding scheme for the case of Gaussian source, interference and side information. This result in Section IV is a fairly straightforward extension of the results in Section II and Section III, but is included for completeness and to make the exposition clear. In Section V, we study the performance of these schemes when the SNR of the channel is different from the designed SNR and in Section VI the distortion exponents of these schemes are analyzed. In Section VII, we consider the problem of transmitting a Gaussian source in the absence of an interference, but when the channel bandwidth is smaller than the source bandwidth and show how the HDA Costa coding scheme is useful. Finally, in Section VIII, we consider the problem of broadcasting a Gaussian source to two users through AWGN channels and propose a joint source channel coding scheme based on HDA Costa coding.

We use the following notation in this paper. Vectors are denoted by bold face letters such as  $\mathbf{x}$ . Upper case letters are used to denote scalar random variables. When considering a sequence of i.i.d random variables, a single upper case letter is used to denote each component of the random vector.

## II. TRANSMISSION OF A GAUSSIAN SOURCE OVER A GAUSSIAN CHANNEL WITH INTERFERENCE KNOWN ONLY AT THE TRANSMITTER

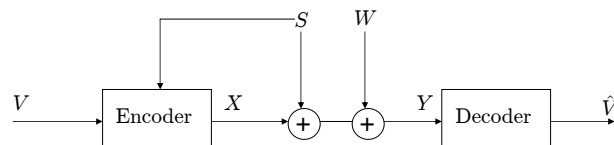


Fig. 1. Block diagram of the joint source channel coding problem with interference known only at the transmitter.

We first consider the problem of transmitting  $N$  samples of a real analog source  $\mathbf{v} \in \mathbb{R}^N$  (this corresponds to  $K = N$ ), with components  $V$  which are independent Gaussian random variables with  $V \sim \mathcal{N}(0, \sigma_v^2)$  in  $N$  uses

of an AWGN channel with noise variance  $\sigma^2$  in the presence of an interference  $\mathbf{s} \in \mathbb{R}^N$  which is known to the transmitter but unknown to the receiver. Further, let us assume that  $S$ 's are a sequence of real i.i.d Gaussian random variables with zero mean and variance  $Q$  and let the input power to the channel  $\mathbb{E}[X^2]$  be constrained to be  $P$ . The problem setup is shown schematically in Fig. 1. The received signal  $\mathbf{y}$  is given by

$$\mathbf{y} = \mathbf{x} + \mathbf{s} + \mathbf{w} \quad (1)$$

where  $\mathbf{s}$  is the interference and  $\mathbf{w}$  is the AWGN.

The optimal distortion of  $\frac{\sigma_v^2}{(1+\frac{P}{\sigma^2})}$  can be obtained even in the presence of the interference by using the following (obvious) separate source and channel coding scheme.

#### A. Separation based scheme with Costa coding (Digital Costa Coding)

We first quantize the source using an optimal quantizer to produce an index  $m \in \{1, 2, \dots, 2^{NR}\}$ , where  $R = \frac{1}{2} \log(1 + \frac{P}{\sigma^2}) - \epsilon$ . Then, the index is transmitted using Costa's writing on dirty paper coding scheme [9]. Since the quantizer output is digital information, we refer to this scheme as digital Costa coding. We briefly review this here to make it easier to describe our proposed techniques later on.

Let  $U$  be an auxiliary random variable given by

$$U = X + \alpha S \quad (2)$$

where  $X \sim \mathcal{N}(0, P)$  is independent of  $S$  and  $\alpha = \frac{P}{P + \sigma^2}$ .

We first create an  $N$ -length i.i.d Gaussian code book  $\mathcal{U}$  with  $2^{N(I(U;Y)-\delta)}$  codewords, where each component of the codeword is Gaussian with zero mean and variance  $P + \alpha^2 Q$ . Then evenly (randomly) distribute these over  $2^{NR}$  bins. For each  $\mathbf{u}$ , let  $i(\mathbf{u})$  be the index of the bin containing  $\mathbf{u}$ . For a given  $m$ , we look for an  $\mathbf{u}$  such that  $i(\mathbf{u}) = m$  and  $(\mathbf{u}, \mathbf{s})$  are jointly typical. Then, we transmit  $\mathbf{x} = \mathbf{u} - \alpha \mathbf{s}$ . Note that since  $(\mathbf{u}, \mathbf{s})$  are jointly typical, from (2), we can see that  $\mathbf{x} \perp \mathbf{s}$  and satisfies the power constraint.

The received sequence  $\mathbf{y}$  is given by

$$\mathbf{y} = \mathbf{x} + \mathbf{s} + \mathbf{w} \quad (3)$$

At the decoder, we look for a  $\mathbf{u}$  that is jointly typical with  $\mathbf{y}$  and declare  $i(\mathbf{u})$  to be the decoded message. Since  $R = \frac{1}{2} \log(1 + \frac{P}{\sigma^2}) - \epsilon$ , the distortion in  $\mathbf{v}$  given by  $D(R)$ , where  $D$  is the distortion rate function. For a Gaussian source and mean squared error distortion  $D(R) = \sigma_v^2 2^{-2R}$  and, hence, the overall distortion can be made to be arbitrarily close to  $\frac{\sigma_v^2}{(1+\frac{P}{\sigma^2})}$  by a proper choice of  $\epsilon$  and  $\delta$ .

While the above scheme is straightforward, in the following three sections we show that there are a few other joint source channel coding schemes, which are also optimal. In fact, there are infinitely many schemes which are optimal. Although, these schemes are all optimal when the channel SNR is known at the transmitter, their performance is in general different when there is an SNR mismatch. The joint source channel coding schemes to be discussed in the next sections have advantages over the separation based scheme discussed in such a situation.

#### B. Hybrid Digital Analog Costa Coding

Let us now describe a joint source-channel coding scheme where the source  $\mathbf{v}$  is not explicitly quantized. We refer to this scheme as hybrid digital analog (HDA) Costa coding for which the code construction, encoding and decoding procedures are as follows.

We first define an auxiliary random variable  $U$  given by

$$U = X + \alpha S + \kappa V \quad (4)$$

where  $X \sim \mathcal{N}(0, P)$  and  $X$ ,  $S$  and  $V$  are pairwise independent.

- 1) Codebook generation: Generate a random i.i.d code book  $\mathcal{U}$  with  $2^{NR_1}$  sequences, where each component of each codeword is Gaussian with zero mean and variance  $P + \alpha^2 Q + \kappa^2 \sigma_v^2$ .
- 2) Encoding: Given an  $\mathbf{s}$  and  $\mathbf{v}$ , find a  $\mathbf{u}$  such that  $(\mathbf{u}, \mathbf{s}, \mathbf{v})$  are jointly typical with respect to the distribution obtained from the model in (4) and transmit  $\mathbf{x} = \mathbf{u} - \alpha \mathbf{s} - \kappa \mathbf{v}$ . If such an  $\mathbf{u}$  cannot be found, we declare an encoder failure. Let  $P_{e_1}$  be the probability of an encoder failure.

From standard arguments on typicality and its extensions to the infinite alphabet case [15], it follows that  $P_{e_1} \rightarrow 0$  as  $N \rightarrow \infty$  provided

$$R_1 > I(U; S, V) \quad (5)$$

$$= h(U) - h(U|S, V) \quad (6)$$

$$= h(U) - h(X|S, V) \quad (7)$$

$$= h(U) - h(X) \quad (8)$$

$$= \frac{1}{2} \log \frac{P + \alpha^2 Q + \kappa^2 \sigma_v^2}{P} \quad (9)$$

where the results follow because  $X = U - \alpha S - \kappa V$  and  $X \perp S, V$ . Notice that when a  $\mathbf{u}$  that is jointly typical with  $\mathbf{s}$  and  $\mathbf{v}$  is found,  $\mathbf{x}$  satisfies the power constraint.

- 3) Decoding : The received signal is  $\mathbf{y} = \mathbf{x} + \mathbf{s} + \mathbf{w}$ . At the decoder, we look for an  $\mathbf{u}$  that is jointly typical with  $\mathbf{y}$ . If such a unique  $\mathbf{u}$  can be found, we declare  $\mathbf{u}$  as the decoder output or, else, we declare a decoder failure. Let  $P_{e_2}$  be the probability of the event that the decoder output is not equal to the encoded  $\mathbf{u}$  (this includes the probability of decoder failure as well as the probability of a decoder error).

In order to analyze  $P_{e_2}$ , consider the equivalent communication channel between  $U$  and  $Y$ . Notice that we have in effect transmitted a codeword  $\mathbf{u}$  from a random i.i.d codebook for  $U$  with  $2^{NR_1}$  codewords through the equivalent channel whose output is  $\mathbf{y}$ . Again, from the extension of joint typicality to the infinite alphabet case,  $P_{e_2} \rightarrow 0$  as  $N \rightarrow \infty$  provided that

$$\begin{aligned} I(U; Y) &> R_1 \\ I(U; Y) &= h(U) - h(U|Y) \\ &= h(U) - h(U - \alpha Y|Y) \\ &= h(U) - h(X + \alpha S + \kappa V - \alpha X - \alpha S - \alpha W|Y) \\ &= h(U) - h(\kappa V + (1 - \alpha)X - \alpha W|Y) \end{aligned} \quad (10)$$

Now, let us choose

$$\alpha = \frac{P}{P + \sigma^2} \quad (11)$$

$$\kappa^2 = \frac{P^2}{(P + \sigma^2)\sigma_v^2} - \frac{\epsilon}{\sigma_v^2} \quad (12)$$

For the above choice of  $\alpha$ , it can be seen that

$$\mathbb{E}[(\kappa V + (1 - \alpha)X - \alpha W)Y] = 0$$

and, hence, (10) reduces to

$$\begin{aligned} I(U; Y) &= h(U) - h(\kappa V + (1 - \alpha)X - \alpha W) \\ &= \frac{1}{2} \log \frac{P + \alpha^2 Q + \kappa^2 \sigma_v^2}{P - \epsilon} \end{aligned} \quad (13)$$

Hence,  $P_{e_2}$  can be made arbitrarily small as long as

$$R_1 < \frac{1}{2} \log \frac{P + \alpha^2 Q + \kappa^2 \sigma_v^2}{P - \epsilon} \quad (14)$$

Combining this with the condition for encoder failure,  $P_{e_1}$  and  $P_{e_2}$  can both be made arbitrarily small provided

$$\frac{1}{2} \log \frac{P + \alpha^2 Q + \kappa^2 \sigma_v^2}{P} < R_1 < \frac{1}{2} \log \frac{P + \alpha^2 Q + \kappa^2 \sigma_v^2}{P - \epsilon} \quad (15)$$

Therefore, by choosing an  $\epsilon_1$ ,  $0 < \epsilon_1 < \epsilon$  and  $R_1 = \frac{1}{2} \log \frac{P + \alpha^2 Q + \kappa^2 \sigma_v^2}{P - \epsilon_1}$  we can satisfy (15) and make  $P_{e_1} \rightarrow 0$  and  $P_{e_2} \rightarrow 0$  as  $N \rightarrow \infty$ .

- 4) Estimation: If there is no decoding failure, we form the final estimate of  $\mathbf{v}$  as an MMSE estimate of  $\mathbf{v}$  from  $[\mathbf{y} \ \mathbf{u}]$ . After some algebra this is given by ,

$$\hat{\mathbf{v}} = \frac{\kappa\sigma_v^2}{P - \epsilon}(\mathbf{u} - \alpha\mathbf{y}) \quad (16)$$

The distortion is then given by,

$$E[(V - \hat{V})^2] = \frac{\sigma_v^2}{1 + \frac{P}{\sigma^2}} \frac{P}{P - \epsilon} \leq \frac{\sigma_v^2}{1 + \frac{P}{\sigma^2}} + \delta(\epsilon) \quad (17)$$

with  $\delta(\epsilon)$  is vanishing for arbitrarily small  $\epsilon$ . If an encoder or decoder failure was declared, we set the estimate of  $\mathbf{v}$  to be the zero vector. However, as shown above the probability of these events can be made arbitrarily small and, hence, they do not contribute to the overall distortion, which can be seen to be arbitrarily close to the optimal distortion achievable in the absence of the interference.

We have presented a joint source channel coding scheme in the presence of an interference known only to the transmitter. The use of the term hybrid digital analog Costa coding needs some explanation. The scheme is not entirely analog in that the auxiliary random variable is from a discrete codebook. However, in contrast to digital Costa coding, the source is not explicitly quantized and is embedded into the transmitted signal  $\mathbf{x}$  in an analog fashion. This is the reason for calling this as HDA Costa coding and this has some interesting consequences which are discussed in the following section.

Another feature of the HDA Costa coding scheme is that it does not make use of binning, rather it needs a single quantizer codebook that is also a good channel code. In practice, this may have some impact on the design since there are ensembles of codes that are provably good quantizers and channel codes [13]. In the Gaussian case, good lattices that are both good for coding and for quantization are known. The binning approach however, requires a nesting condition. That is, the fine code must be a good channel code, but it must contain a subcode (or a coarse code) and its cosets that must be good quantizers. This may be a more difficult condition to obtain in practice.

### C. Superimposed digital and HDA Costa coding scheme

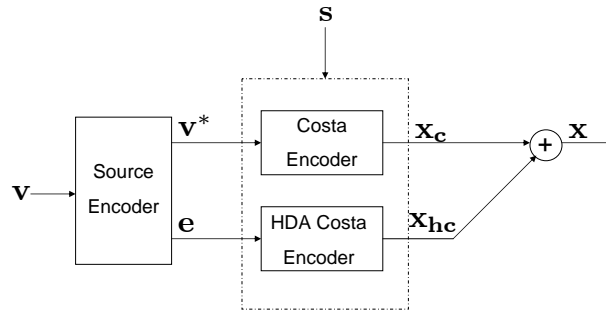


Fig. 2. Encoder model for superimposed coding

Recently in [6], Bross, Lapidoth and Tinguely considered the problem of transmitting  $N$  samples of a Gaussian source in  $N$  uses of an AWGN channel, in the absence of the interferer. They showed that there are infinitely many superposition based schemes, which contain pure separation based scheme and uncoded transmission as special cases. In this section, we show that the same is true in the presence of an interference also and show the corresponding scheme, which is given in Fig. 2.

The transmitted signal is a superposition of two signals  $\mathbf{x}_c$  and  $\mathbf{x}_{hc}$ , which are the outputs of a digital Costa encoder and an HDA encoder, respectively.

The source is first quantized at a rate of  $R < C$  using an optimal source code and let the quantization error be  $\mathbf{e} = \mathbf{v} - \mathbf{v}^*$ , where  $\mathbf{v}^*$  is the reconstruction. The quantization error  $\mathbf{e}$  has a variance  $\sigma_e^2 = \sigma_v^2 2^{-2R}$ . The first stream in Fig. 2 is a digital Costa encoder that encodes the quantization index by treating  $\mathbf{s}$  as interference and produces the signal  $\mathbf{x}_c$ , which has a power of  $P_C$ . The second stream is a HDA Costa encoder of rate  $R$  which

treats  $\mathbf{s}$  and  $\mathbf{x}_c$  as interference and produces  $\mathbf{x}_{hc}$ , which has a power of  $P_{HC} = P - P_C$ . The transmitted signal is the superposition (sum) of  $\mathbf{x}_c$  and  $\mathbf{x}_{hc}$ .

In the digital Costa encoder in the first stream, the auxiliary random variable is given by  $\mathbf{u}_c = \mathbf{x}_c + \alpha_c \mathbf{s}$  with  $\mathbf{x}_c \perp \mathbf{s}$ . A power of  $P_C = (P + \sigma^2)(1 - 2^{-2R})$  is used in the first stream and  $\alpha_c$  is chosen as  $\frac{P_C}{P_C + P_{HC} + \sigma^2}$ . Note that this corresponds to treating  $\mathbf{x}_{hc}$  as noise in addition to the channel noise.

In the second stream, the quantization error  $\mathbf{e}$  is encoded using an HDA Costa coding scheme and a power of  $P_{HC} = P - P_C = (P + \sigma^2)2^{-2R} - \sigma^2$  is used. Note that since  $R < C$ , the power  $P_{HC}$  is always positive. The auxiliary random variable is chosen as  $\mathbf{u}_{hc} = \mathbf{x}_{hc} + \alpha_{hc}(\mathbf{x}_c + \mathbf{s}) + \kappa \mathbf{e}$ , where  $\mathbf{x}_c + \mathbf{s}$  acts as the net interference. Hence,  $\mathbf{x}_{hc}$  is chosen to be independent of  $\mathbf{x}_c$ ,  $\mathbf{s}$  and  $\mathbf{e}$ , and  $\alpha_{hc}$  is chosen to be  $\frac{P_{HC}}{P_{HC} + \sigma^2}$ .  $\kappa$  is chosen similar to (12) which gives  $\kappa^2 = \frac{P_{HC}^2}{(P_{HC} + \sigma^2)\sigma_v^2 2^{-2R}} - \frac{\epsilon}{\sigma_v^2 2^{-2R}}$ .

At the decoder the quantization index from the first stream is first decoded and the reconstruction  $\mathbf{v}^*$  is obtained. Then, an estimate of the quantization error  $\mathbf{e}$  is obtained from the second stream using the HDA Costa decoder. The overall distortion is the distortion in estimating  $\mathbf{e}$ . Using the analysis of the HDA Costa scheme in Section II-B, this can be seen to be

$$D = \frac{\sigma_e^2}{1 + \frac{(P + \sigma^2)2^{-2R} - \sigma^2}{\sigma^2}} + \delta(\epsilon) = \frac{\sigma_v^2}{1 + \frac{P}{\sigma^2}} + \delta(\epsilon) \quad (18)$$

By choosing  $\epsilon$  to be arbitrarily small we can make  $\delta(\epsilon) \rightarrow 0$  and achieve a distortion of  $D = \frac{\sigma_v^2}{1 + \frac{P}{\sigma^2}}$ , which is the optimal distortion.

Note that for any source coding rate chosen in the first stream namely  $R$ , the resulting distortion is optimal. By varying  $R$ , we can get an infinite family of optimal joint source channel coding schemes.

#### D. Generalized Hybrid Costa coding

In the previous section, we described a superposition technique. In this section we show a scheme that does not explicitly do superposition. Moreover this also introduces an interesting scheme that is intermediate between HDA Costa having no bins to the digital Costa having bins corresponding to the capacity of the channel.

Once again we quantize the source  $\mathbf{v}$  to  $\mathbf{v}^*$  at a rate  $R$ , that is strictly lesser than the channel capacity, using an optimal vector quantizer. Let  $\mathbf{e} = \mathbf{v} - \mathbf{v}^*$  be the quantization error vector. Note that for an optimal quantizer, as the Rate-Distortion limit is approached, the quantization error  $\mathbf{e}$  will be Gaussian.

We next define an auxiliary random variable  $U$  given by

$$U = X + \alpha S + \kappa_1 E \quad (19)$$

where  $X \sim \mathcal{N}(0, P)$ ,  $E \sim \mathcal{N}(0, \sigma_v^2 2^{-2R})$ , and  $X$ ,  $S$  and  $E$  are independent of each other.  $\alpha$  and  $\kappa_1$  are constants, the choice of which is discussed below

- 1) Codebook generation: Generate a random i.i.d code book  $\mathcal{U}$  with  $2^{NI(U;Y)}$  sequences, where each component of each codeword is Gaussian with zero mean and variance  $P + \alpha^2 Q + \kappa_1^2 \sigma_v^2 2^{-2R}$ . These codewords are uniformly distributed in  $2^{NR}$  bins and this is shared between the encoder and the decoder.
- 2) Encoding: Let  $m$  be the quantization index corresponding to the quantized source  $\mathbf{v}^*$ . Let  $i(\mathbf{u})$  represent the index of a bin that contains  $\mathbf{u}$ . For a given  $m$  find an  $\mathbf{u}$  such that  $i(\mathbf{u}) = m$  and  $(\mathbf{u}, \mathbf{s}, \mathbf{e})$  are jointly typical with respect to the distribution in model (19). We next transmit the vector  $\mathbf{x} = \mathbf{u} - \alpha \mathbf{s} - \kappa_1 \mathbf{e}$ . Note that since  $(\mathbf{u}, \mathbf{s}, \mathbf{e})$  are jointly typical, from (19), we can see that  $\mathbf{x} \perp \mathbf{s}, \mathbf{e}$  and satisfies the power constraint.
- 3) Decoding : The received signal is  $\mathbf{y} = \mathbf{x} + \mathbf{s} + \mathbf{w}$ . At the decoder, we look for an  $\mathbf{u}$  that is jointly typical with  $\mathbf{y}$ . If such a unique  $\mathbf{u}$  can be found, we declare  $\mathbf{u}$  as the decoder output or, else, we declare a decoder failure. Next we make an estimate of  $\mathbf{e}$  from  $\mathbf{u}$  and  $\mathbf{y}$ .

We can see by similar Gelfand-Pinsker coding arguments that  $R < I(U; Y) - I(U; S, E)$ . Note

$$\begin{aligned}
I(U; Y) &= I(U; S, E) \\
&= h(U|S, E) - h(U|Y) \\
&= h(X) - h(U - \alpha Y|Y) \\
&= h(X) - h(\kappa_1 E + (1 - \alpha)X - \alpha W|Y) \\
&\stackrel{(a)}{=} h(X) - h(\kappa_1 E + (1 - \alpha)X - \alpha W) \\
&= \frac{1}{2} \log \left( \frac{P}{\kappa_1^2 \sigma_v^2 2^{-2R} + (1 - \alpha)^2 P + \alpha^2 \sigma^2} \right) \\
&\stackrel{(b)}{>} R
\end{aligned} \tag{20}$$

In (20) we choose  $\alpha = \frac{P}{P + \sigma^2}$  and  $\kappa_1^2 = \frac{P}{P + \sigma^2} \frac{(P + \sigma^2) - \sigma^2 2^{2R}}{\sigma_v^2} - \frac{\epsilon((P + \sigma^2) - \sigma^2 2^{2R})}{P \sigma_v^2}$ . The choice of  $\alpha$  ensures  $(1 - \alpha)X - \alpha W$  is orthogonal to  $Y$  to get the equality in (a).  $\kappa_1$  is chosen as above to satisfy the inequality in (b). This shows that we can decode the codeword  $\mathbf{u}$  with a very high probability and we can decode the message  $m = i(\mathbf{u})$  and  $\mathbf{v}^*$ .

- 4) Estimation: If there is no decoding failure, we form the final estimate of  $\mathbf{v}$  as an MMSE estimate of  $\mathbf{v}$  from  $[\mathbf{v}^* \mathbf{u} \mathbf{y}]$ . The estimate can be obtained as follows. Let us define  $\sigma_e^2 = \sigma_v^2 2^{-2R}$ . Let  $\mathbf{\Lambda}$  be the covariance matrix of  $[V^* \ U \ Y]^T$  and let  $\mathbf{\Gamma}$  be the correlation vector between  $V$  and  $[V^* \ U \ Y]^T$ . Then,  $\mathbf{\Lambda}$  and  $\mathbf{\Gamma}$  are given by

$$\mathbf{\Lambda} = \begin{pmatrix} \sigma_v^2 - \sigma_e^2 & 0 & 0 \\ 0 & P + \kappa_1^2 \sigma_e^2 + \alpha^2 Q & P + \alpha Q \\ 0 & P + \alpha Q & P + Q + \sigma^2 \end{pmatrix} \text{ and } \mathbf{\Gamma} = (\sigma_v^2 - \sigma_e^2 \quad \kappa_1 \sigma_e^2 \quad 0)^T$$

The coefficients of the linear MMSE estimate are given by  $\mathbf{\Lambda}^{-1} \mathbf{\Gamma}$  and the minimum mean-squared error is given by

$$D = \sigma_v^2 - \mathbf{\Gamma}^T \mathbf{\Lambda}^{-1} \mathbf{\Gamma} = \left( \frac{\sigma_v^2}{1 + \frac{P}{\sigma^2}} \right) + \delta(\epsilon)$$

where  $\delta(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ . Thus, in the limit of  $\epsilon \rightarrow 0$ ,  $D = \left( \frac{\sigma_v^2}{1 + \frac{P}{\sigma^2}} \right)$ .

It must be noted that this scheme is an intermediate between digital Costa coding scheme with the maximum possible bins equal to the capacity of the channel and the analog Costa coding scheme with no bins. Thus we can get a family of schemes with varying bins for the Gaussian channel.

The generalized hybrid Costa coding scheme appears to be closely related to the superimposed digital and HDA Costa coding schemes. The subtle difference however is in the generalized hybrid Costa coding scheme, the transmitted signal  $X$  is not a superposition of two streams as seen in the superposition case.

### III. TRANSMISSION OF A GAUSSIAN SOURCE THROUGH A CHANNEL WITH SIDE INFORMATION AVAILABLE ONLY AT THE RECEIVER

In this section we consider the problem of transmitting a discrete-time analog source over a Gaussian noise channel when the receiver has some side information about the source. This problem is a dual of the problem considered in the previous section and is considered here for the sake of completeness. Consider the system model as follows. Let  $\mathbf{v} \in \mathbb{R}^N$  be the discrete-time analog source where  $V$ 's are independent Gaussian random variables with  $V \sim \mathcal{N}(0, \sigma_v^2)$ . Let  $\mathbf{v}' \in \mathbb{R}^N$  be the side information that is known only at the receiver. The correlation between the source and the side information is modeled as

$$V = V' + Z \tag{21}$$

where  $Z \sim \mathcal{N}(0, \sigma_z^2)$  and  $V'$  is i.i.d Gaussian. Here  $V'$  and  $Z$  are mutually independent random variables. The source  $\mathbf{v}$  must be encoded into  $\mathbf{x}$  and transmitted over an AWGN channel and the received signal is

$$\mathbf{y} = \mathbf{x} + \mathbf{w} \tag{22}$$

where  $\mathbf{x}$  satisfies a power constraint  $P$  and  $\mathbf{w}$  is AWGN having a noise variance of  $\sigma^2$ . The following schemes can be shown to be optimal for this case.

#### A. Separation Based Scheme with Wyner Ziv Coding (Digital Wyner Ziv Coding)

One strategy is using a separation scheme with an optimal Wyner-Ziv code of rate  $R$  followed by a channel code. We also refer to this scheme as the digital Wyner-Ziv scheme. We briefly explain the digital Wyner-Ziv scheme and then establish our information theoretic model for the HDA Wyner-Ziv coding scheme.

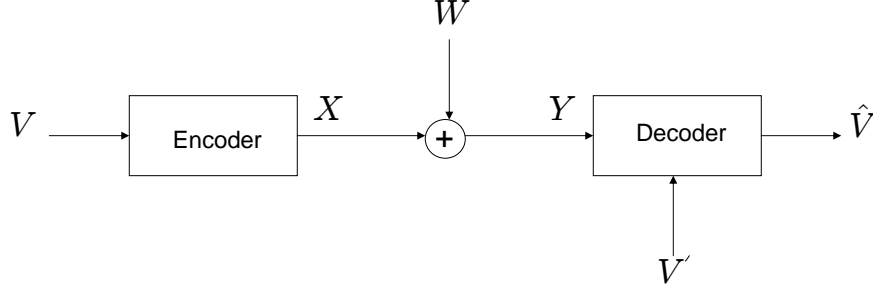


Fig. 3. Block diagram of the joint source channel coding problem with side information known only at the receiver.

Suppose the side information is available both at the encoder as well as the receiver, the best possible distortion is  $D = \frac{\sigma_z^2}{1 + \frac{P}{\sigma_z^2}}$ . The same distortion can be achieved using the following scheme and is a direct consequence of Wyner and Ziv's result [16]. This can be achieved as follows,

Let  $U$  be an auxiliary random variable given by

$$U = \sqrt{\alpha}V + B \quad (23)$$

where  $\alpha = 1 - \frac{D}{\sigma_z^2} = \frac{P}{P + \sigma_z^2}$  and  $B \sim \mathcal{N}(0, D)$ . We create an  $N$ -length i.i.d Gaussian code book  $\mathcal{U}$  with  $2^{NI(U;V)}$  codewords, where each component of the codeword is Gaussian with zero mean and variance  $\alpha\sigma_v^2 + D$  and evenly distribute them over  $2^{NR}$  bins. Let  $i(\mathbf{u})$  be the index of the bin containing  $\mathbf{u}$ . For each  $\mathbf{v}$ , find an  $\mathbf{u}$  such that  $(\mathbf{u}, \mathbf{v})$  are jointly typical. The index  $i(\mathbf{u})$  is the Wyner-Ziv source coded index. The index  $i(\mathbf{u})$  is encoded using an optimal channel code of rate arbitrarily close to  $\frac{1}{2} \log(1 + \frac{P}{\sigma_z^2})$  and transmitted over the channel. At the receiver decoding of the index  $i(\mathbf{u})$  is possible with high probability as an optimal code book for the channel is used. Next for the decoded  $i(\mathbf{u})$  we look for an  $\mathbf{u}$  in the bin whose index is  $i(\mathbf{u})$  such that  $(\mathbf{u}, \mathbf{v}')$  are jointly typical. From  $\mathbf{v}'$  and the decoded  $\mathbf{u}$  we make an estimate of the source  $\mathbf{v}$  as follows.

$$\hat{\mathbf{v}} = \mathbf{v}' + \sqrt{\alpha}(\mathbf{u} - \sqrt{\alpha}\mathbf{v}') \quad (24)$$

This yields the optimal distortion  $D$ .

#### B. Hybrid Digital Analog Wyner Ziv Coding

In this section, we discuss a different joint source channel coding scheme that does not involve quantizing the source explicitly. This scheme is quite similar to the modulo lattice modulation scheme in [2]; the difference being that a nested lattice is not used. The auxiliary random variable  $U$  is generated as follows.

$$U = X + \kappa V \quad (25)$$

where  $\kappa$  is defined as  $\kappa^2 = \frac{P^2}{(P + \sigma_z^2)\sigma_z^2} - \frac{\epsilon}{\sigma_z^2}$  and  $X \sim \mathcal{N}(0, P)$ .

- 1) Codebook generation: Generate a random i.i.d code book  $\mathcal{U}$  with  $2^{NR_1}$  sequences, where each component of each codeword is Gaussian with zero mean and variance  $P + \kappa^2\sigma_v^2$ . This codebook is shared between the encoder and the decoder.
- 2) Encoding: For a given  $\mathbf{v}$  find an  $\mathbf{u}$  such that  $(\mathbf{u}, \mathbf{v})$  are jointly typical and transmit  $\mathbf{x} = \mathbf{u} - \kappa\mathbf{v}$ . This is possible with arbitrarily high probability if  $R_1 > I(U; V)$



- 3) Decoding: The received signal is  $\mathbf{y} = \mathbf{x} + \mathbf{w}$ . Find an  $\mathbf{u}$  such that  $(\mathbf{v}', \mathbf{y}, \mathbf{u})$  are jointly typical. A unique such  $\mathbf{u}$  can be found with arbitrarily high probability if  $R_1 < I(U; V', Y)$ . We next show below that we can choose an  $R_1$  to satisfy  $I(U; V) < R_1 < I(U; V', Y)$ . This requires  $I(U; V) < I(U; V', Y)$  which can be shown as follows

$$\begin{aligned}
I(U; V', Y) &= h(U) - h(U|V', Y) \\
&= h(U) - h(U - \kappa V' - \alpha Y|V', Y) \\
&= h(U) - h(\kappa Z + (1 - \alpha)X - \alpha W|V', Y) \\
&\stackrel{(a)}{=} h(U) - h(\kappa Z + (1 - \alpha)X - \alpha W) \\
&= \frac{1}{2} \log \left( \frac{P + \kappa^2 \sigma_v^2}{\kappa^2 \sigma_z^2 + (1 - \alpha)^2 P + \alpha^2 \sigma^2} \right) \\
&\stackrel{(b)}{=} \frac{1}{2} \log \left( \frac{P + \kappa^2 \sigma_v^2}{P} \right) + \delta(\epsilon) \\
&= h(U) - h(U|V) + \delta(\epsilon) \\
&= I(U; V) + \delta(\epsilon)
\end{aligned} \tag{26}$$

In (26), (a) follows because  $(\kappa Z + (1 - \alpha)X - \alpha W)$  is independent of  $Y$  and  $V'$ . (b) follows because we can always find a  $\delta(\epsilon) > 0$  for the choice of  $\kappa^2 = \frac{P^2}{(P + \sigma^2)\sigma_z^2} - \frac{\epsilon}{\sigma_z^2}$ . Hence from knowing  $\mathbf{v}'$ ,  $\mathbf{u}$  and  $\mathbf{y}$  we can make an estimate of  $\mathbf{v}$ . Since all random variables are Gaussian, the optimal estimate is a linear MMSE estimate which can be computed as follows.

Let  $\mathbf{\Lambda}$  be the covariance matrix of  $[V' \ U \ Y]^T$  and let  $\mathbf{\Gamma}$  be the correlation between  $V$  and  $[V' \ U \ Y]^T$ .  $\mathbf{\Lambda}$  and  $\mathbf{\Gamma}$  are given by

$$\mathbf{\Lambda} = \begin{pmatrix} \sigma_v^2 - \sigma_z^2 & \kappa(\sigma_v^2 - \sigma_z^2) & 0 \\ \kappa(\sigma_v^2 - \sigma_z^2) & P + \kappa^2 \sigma_v^2 & P \\ 0 & P & P + \sigma^2 \end{pmatrix} \text{ and } \mathbf{\Gamma} = \begin{pmatrix} \sigma_v^2 - \sigma_z^2 & \kappa \sigma_v^2 & 0 \end{pmatrix}^T$$

The coefficients of the linear MMSE estimate are given by  $\mathbf{\Lambda}^{-1} \mathbf{\Gamma}$  and this yields the optimal MMSE estimate which is given below as,

$$\hat{\mathbf{v}} = \mathbf{v}' + \frac{\kappa \sigma_z^2}{P} (\mathbf{u} - \kappa \mathbf{v}' - \alpha \mathbf{y}) \tag{27}$$

The distortion  $D$  is given by

$$\begin{aligned}
D &= E[(\mathbf{v} - \hat{\mathbf{v}})^2] \\
&= E[(\mathbf{v} - \mathbf{v}' - \frac{\kappa \sigma_z^2}{P} (\mathbf{u} - \kappa \mathbf{v}' - \alpha \mathbf{y}))^2] \\
&= E[(\mathbf{z} - \frac{\kappa \sigma_z^2}{P} (\kappa \mathbf{z} + \mathbf{x} - \alpha \mathbf{y}))^2] \\
&= E[(\left(1 - \frac{\kappa^2 \sigma_z^2}{P}\right) \mathbf{z} - \frac{\kappa \sigma_z^2}{P} ((1 - \alpha) \mathbf{x} - \alpha \mathbf{w}))^2] \\
&\stackrel{(a)}{=} \frac{\sigma_z^2}{1 + \frac{P}{\sigma^2}} + \delta(\epsilon)
\end{aligned} \tag{28}$$

Here, (a) follows by using the appropriate values of  $\kappa$  and  $\alpha$ . We once again obtain the optimal distortion  $D$  by making  $\epsilon$  arbitrarily small and  $\delta(\epsilon) \rightarrow 0$ .

It is instructive to compare the performance of this scheme with that of the following naive scheme that would be optimal in the absence of side-information at the receiver. In the naive scheme, the  $\mathbf{v}$  is transmitted directly (analog

transmission). At the receiver, an MMSE estimate of  $\mathbf{v}$  is formed from the received signal  $\mathbf{y}$  and the available side information  $\mathbf{v}'$ . The distortion for this naive scheme can be seen to be  $D_{naive} = \sigma_z^2 / (1 + (P/\sigma^2)\sigma_z^2)$ .

Notice that  $\frac{\partial D_{naive}}{\partial \sigma_z^2}|_{\sigma_z^2=0} = 1$ , whereas for the Wyner-Ziv scheme,  $\frac{\partial D}{\partial \sigma_z^2}|_{\sigma_z^2=0} = \frac{1}{1+P/\sigma^2} < 1$ . At  $\sigma_z^2 = 0$ , both  $D_{naive}$  and  $D$  are zero. i.e. the optimal scheme and the naive scheme approach zero distortion with different slopes.

### C. Superimposed digital and HDA Wyner-Ziv scheme

The above results could also be extended to a form of superimposed digital and analog coding. This is similar to the superimposed digital and HDA Costa coding case discussed in section II-C. We once again have two streams as shown in Fig 4. The first stream uses a rate  $R$  Wyner Ziv code to quantize the source assuming the side information  $\mathbf{v}'$  is known at the receiver and the discrete index is encoded using an optimal channel code to produce the codeword  $\mathbf{x}_1$ . The power allocated to this stream is  $P_{WZ} = (P + \sigma^2)(1 - 2^{-2R})$ . The second stream uses the HDA Wyner-Ziv scheme and produces the output  $\mathbf{x}_2$ . The auxiliary random variable of the HDA scheme is given by

$$U = \kappa_1 V + X_2 \quad (29)$$

with  $X_2 \sim \mathcal{N}(0, P_{HWZ})$ , where  $P_{HWZ} = (P + \sigma^2)2^{-2R} - \sigma^2$  and  $X_2$  and  $V$  are independent. We also choose  $\kappa_1^2 = \frac{P_{HWZ}^2}{(P_{HWZ} + \sigma^2)\sigma_e^2} - \frac{\epsilon}{\sigma_e^2}$  where  $\sigma_e^2 = \sigma_z^2 2^{-2R}$ .

The two streams ( $\mathbf{x}_1$  and  $\mathbf{x}_2$ ) are superimposed and transmitted through the channel. The received signal is given by  $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{w}$ . At the receiver  $\mathbf{x}_1$  is decoded assuming  $\mathbf{x}_2 + \mathbf{w}$  as independent noise and this gives the Wyner-Ziv encoded bits (index). This along with the side information  $\mathbf{v}'$  can be used to make an estimate of the source  $\mathbf{v}$  and we call the estimate as  $\tilde{\mathbf{v}}$ . The random variables corresponding to  $\mathbf{v}$  and  $\tilde{\mathbf{v}}$  are related as

$$V = \tilde{V} + \tilde{Z} \quad (30)$$

with  $\tilde{Z}$  having a variance  $\sigma_z^2 2^{-2R}$ . When the digital part is first decoded and canceled from the received signal, we get an equivalent channel for the HDA Wyner-Ziv scheme with power constraint  $P_{HWZ}$  and channel noise  $\sigma^2$ . We next make a final estimate of  $\mathbf{v}$  using a HDA Wyner-Ziv decoder from the new side information  $\tilde{\mathbf{v}}$ , the observed equivalent channel ( $\mathbf{y} - \mathbf{x}_1$ ) and the decoded  $\mathbf{u}$ . Notice that since the choice of  $\kappa_1^2 = \frac{P_{HWZ}^2}{(P_{HWZ} + \sigma^2)\sigma_e^2} - \frac{\epsilon}{\sigma_e^2}$  where  $\sigma_e^2 = \sigma_z^2 2^{-2R}$  is designed for the side information  $\tilde{\mathbf{v}}$ , this ensures decoding of  $\mathbf{u}$  with arbitrarily high probability. The achievable distortion is then given as follows.

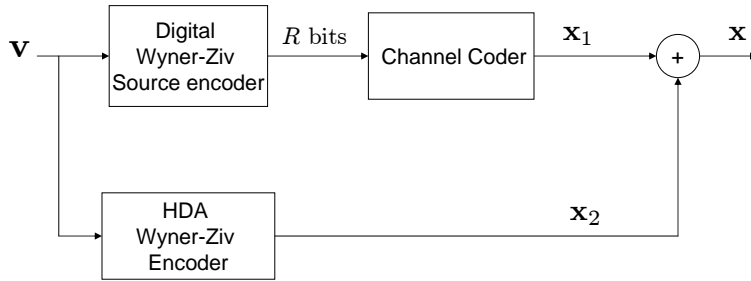


Fig. 4. Block diagram of the encoder of the superimposed digital and HDA Wyner-Ziv scheme.

$$\begin{aligned}
 D &\stackrel{(a)}{=} \sigma_e^2 \frac{\sigma^2}{P_{HWZ} + \sigma^2} + \delta(\epsilon) \\
 &= \sigma_z^2 2^{-2R} \frac{\sigma^2}{P_{HWZ} + \sigma^2} + \delta(\epsilon) \\
 &\stackrel{(b)}{=} \sigma_z^2 \frac{P_{HWZ} + \sigma^2}{P + \sigma^2} \frac{\sigma^2}{P_{HWZ} + \sigma^2} + \delta(\epsilon) \\
 &= \frac{\sigma_z^2}{1 + \frac{P}{\sigma^2}} + \delta(\epsilon)
 \end{aligned} \quad (31)$$

Here in (31) (a) follows since we assume that the first stream is decoded with high probability and apply the results of HDA Wyner-Ziv decoding with the side information  $\tilde{\mathbf{v}}$ . Also (b) follows since  $P_{HWZ} = (P + \sigma^2)2^{-2R} - \sigma^2$ . The optimal distortion  $\frac{\sigma_z^2}{1 + \frac{P}{\sigma_z^2}}$  can be obtained by making  $\epsilon$  arbitrarily small and  $\delta(\epsilon) \rightarrow 0$ . Notice that for any rate  $R$ ,  $0 \leq R < C$ , where  $C$  is the capacity of the AWGN channel, there is a corresponding power allocation for  $P_{HWZ} = (P + \sigma^2)2^{-2R} - \sigma^2$  for which the overall scheme is optimal. Thus, there are infinitely many schemes which are optimal with the digital Wyner-Ziv corresponding to  $P_{HWZ} = 0$  and the HDA Wyner-Ziv corresponding to  $P_{HWZ} = P$  and  $R = 0$ .

Further, we would like to mention that there is another way to get a family of optimal schemes using the HDA Wyner-Ziv scheme. Here, the source  $\mathbf{v}$  is encoded using a HDA Wyner-Ziv encoder to the sequence  $\mathbf{x}$ . The auxiliary random variable  $U$  is given by

$$U = \kappa V + X \quad (32)$$

where  $\kappa^2 = \frac{P^2}{(P + \sigma^2)\sigma_v^2} - \frac{\epsilon}{\sigma_v^2}$ . The sequence  $\mathbf{x}$  can be treated as an i.i.d Gaussian source and, hence, the family of schemes proposed by Bross, Lapidot and Tinguely [6] can be applied on  $\mathbf{x}$ . The scheme proposed in [6] quantizes the analog source, which in this case is  $\mathbf{x}$  to a quantization index and is sent over the Gaussian channel along with the uncoded analog source (here  $\mathbf{x}$ ) with the appropriate power scaling. At the receiver we can obtain an optimal estimate of  $\mathbf{x}$  by first decoding the quantized index and then making an estimate on the analog source. Notice that the HDA Wyner-Ziv receiver only requires an optimal MMSE estimate of  $\mathbf{x}$ , which can be obtained using the family of schemes in [6]. Hence the resulting distortion in  $\mathbf{v}$  is still optimal. To establish this claim we need to show that  $\mathbf{u}$  can be decoded with arbitrarily high probability and an optimal estimate of  $\mathbf{v}$  must be made using  $\mathbf{u}$  and the MMSE estimate  $\hat{\mathbf{x}}$ .

We next show below that  $I(U; V) < I(U; V', \hat{X})$ . Hence, we can choose a codebook for  $\mathbf{u}$  with  $2^{nR_1}$  codewords such that  $I(U; V) < R_1 < I(U; V', \hat{X})$ . Since  $I(U; V) < R_1$ , we can find a  $\mathbf{u}$  that is jointly typical with  $\mathbf{v}$  with probability close to 1 and since  $R_1 < I(U; V', \hat{X})$ ,  $\mathbf{u}$  can be decoded with high probability from  $(V', \hat{\mathbf{x}})$ .

$$\begin{aligned} I(U; V', \hat{X}) &= h(U) - h(U|V', \hat{X}) \\ &= h(U) - h(U - \kappa V' - \hat{X}|V', \hat{X}) \\ &= h(U) - h(\kappa Z + X - \hat{X}|\hat{X}, V') \\ &\stackrel{(a)}{=} h(U) - h(\kappa Z + X - \hat{X}) \\ &\stackrel{(b)}{=} \frac{1}{2} \log \left( \frac{P + \kappa^2 \sigma_v^2}{\kappa^2 \sigma_z^2 + \alpha \sigma^2} \right) \\ &= \frac{1}{2} \log \left( \frac{P + \kappa^2 \sigma_v^2}{P} \right) + \delta(\epsilon) \\ &= h(U) - h(U|V) + \delta(\epsilon) \\ &= I(U; V) + \delta(\epsilon) \end{aligned} \quad (33)$$

In (33), (a) follows because  $(X - \hat{X})$  is orthogonal to  $\hat{X}$  and hence  $(\kappa Z + X - \hat{X})$  is independent of  $\hat{X}$  and  $V'$ , (b) follows because  $X - \hat{X}$  is Gaussian with variance  $\alpha \sigma^2$  and is orthogonal to  $Z$ . The estimate of  $\mathbf{v}$  is then given by

$$\hat{\mathbf{v}} = \mathbf{v}' + \frac{\kappa \sigma_z^2}{P} (\mathbf{u} - \kappa \mathbf{v}' - \hat{\mathbf{x}}) \quad (34)$$

The resulting distortion can be obtained by following the steps similar to those in (28) which can be found to be optimal.

#### IV. TRANSMISSION OF A GAUSSIAN SOURCE WITH INTERFERENCE AT THE TRANSMITTER AND SIDE INFORMATION AT THE RECEIVER

In this section, we consider the problem of transmitting a Gaussian source  $\mathbf{v}$  through an AWGN channel with channel noise variance  $\sigma^2$  in the presence of an interference  $\mathbf{s}$  known only at the transmitter and in the presence

of side information  $\mathbf{v}'$  known only at the receiver. The side information  $\mathbf{v}'$  is assumed to be related to the source  $\mathbf{v}$  according to

$$V = V' + Z$$

where  $Z \sim \mathcal{N}(0, \sigma_z^2)$  and is independent of  $V'$ .

A similar model has been considered by Merhav and Shamai [11] for a more general setup where the source and side information are not assumed to be Gaussian. They show that a separation based approach of Wyner-Ziv coding followed by Gelfand-Pinsker coding is optimal. Here, we propose a joint-source channel coding scheme when the source and channel noise are Gaussian. The proposed scheme is easily obtained by combining the results from the previous two sections. It must be noted that a similar joint source channel coding scheme using nested lattices and dither has been shown in [7]. However, our scheme is based only on random code books.

To establish our scheme we can combine the results from the previous two sections as follows. Choose the auxiliary random variable  $U$  such that

$$U = X + \alpha S + \kappa V \quad (35)$$

with  $\kappa^2 = \frac{P^2}{(P+\sigma^2)\sigma_z^2} - \frac{\epsilon}{\sigma_z^2}$  and  $\alpha = \frac{P}{P+\sigma^2}$ . Further, let  $X \sim \mathcal{N}(0, P)$ ,  $S \sim \mathcal{N}(0, Q)$  and  $V \sim \mathcal{N}(0, \sigma_v^2)$  and let  $X$ ,  $S$  and  $V$  be pairwise independent. A codebook  $\mathcal{U}$  is obtained by generating  $2^{nR_1}$  code sequences for  $\mathbf{u}$  and this is shared between the encoder and decoder. At the encoder, the source  $\mathbf{v}$  is encoded by choosing an  $\mathbf{x}$  that is jointly typical with  $\mathbf{u}, \mathbf{v}$  and  $\mathbf{s}$ . Such a  $\mathbf{u}$  exists with high probability if we have chosen  $R_1 > I(U; S, V)$ . Now  $\mathbf{x}$  is transmitted over the channel. The received signal vector  $\mathbf{y}$  is given as

$$\mathbf{y} = \mathbf{x} + \mathbf{s} + \mathbf{w}$$

At the decoder,  $\mathbf{u}$  is decoded by looking for a  $\mathbf{u}$  that is jointly typical with  $\mathbf{y}$  and the side information  $\mathbf{v}'$ . Using standard arguments on joint-typicality, it can be seen that a unique such  $\mathbf{u}$  exists with high probability if  $R_1 < I(U; Y, V')$ . We now show that  $I(U; S, V) < I(U; Y, V')$ . This implies that there exists an  $R_1$ , such that  $I(U; S, V) < R_1 < I(U; Y, V')$  which satisfies the requirements at the encoder and the decoder.

$$\begin{aligned} I(U; Y, V') &= h(U) - h(U|Y, V') \\ &= h(U) - h(U - \alpha Y - \kappa V'|Y, V') \\ &= h(U) - h(\kappa Z + (1 - \alpha)X - \alpha W|Y, V') \\ &\stackrel{(a)}{=} h(U) - h(\kappa Z + (1 - \alpha)X - \alpha W) \\ &= \frac{1}{2} \log \left( \frac{P + \alpha^2 Q + \kappa^2 \sigma_v^2}{P} \right) + \delta(\epsilon) \\ &= I(U; S, V) + \delta(\epsilon) \end{aligned} \quad (36)$$

where (a) follows since  $\kappa Z + (1 - \alpha)X - \alpha W$  is orthogonal to  $Y$  and  $V'$ . Then an optimal linear MMSE estimate of  $\mathbf{v}$  is formed from the side information  $\mathbf{v}'$ , the received vector  $\mathbf{y}$  and the vector  $\mathbf{u}$ . By using the argument as in section. III-B, the MMSE estimate is given by

$$\hat{\mathbf{v}} = \mathbf{v}' + \frac{\kappa \sigma_z^2}{P} (\mathbf{u} - \kappa \mathbf{v}' - \alpha \mathbf{y}) \quad (37)$$

The resulting distortion can be obtained by following steps similar to (28) and can be seen to be  $D = \frac{\sigma_z^2}{1 + \frac{P}{\sigma_z^2}}$ , which is the optimal distortion.

## V. ANALYSIS OF THE SCHEMES FOR SNR MISMATCH

In this section, we consider the performance of the above JSCC schemes for the case of SNR mismatch where we design the scheme to be optimal for a channel noise variance of  $\sigma^2$ , but the actual noise variance is  $\sigma_a^2$ .

Separation based digital schemes suffer from a pronounced threshold effect. When the channel SNR is worse than the designed SNR, the index cannot be decoded and when the channel SNR is better than the designed SNR, the distortion is limited by the quantization and does not improve. However, the hybrid digital analog schemes considered offer better performance in this situation.

Let us consider the joint source channel coding setup with side information at both the transmitter and receiver and  $\sigma_a^2 < \sigma^2$ . We can decode  $\mathbf{u}$  at the receiver when the SNR is better than the designed SNR and make an estimate of the source from the various observations at the receiver as shown below.

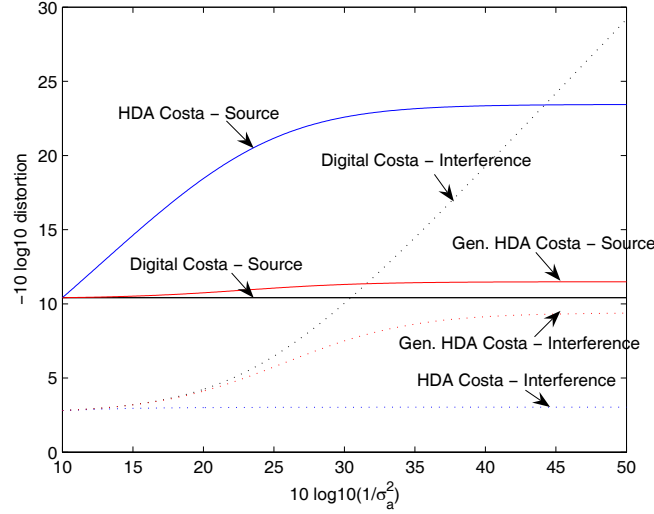


Fig. 5. Performance of the different Costa coding schemes for the joint source channel coding problem.

$$U = X + \alpha S + \kappa_w V \quad (38)$$

$$V = V' + Z \quad (39)$$

$$Y = X + S + W_a \quad (40)$$

where  $\kappa_w = \sqrt{\frac{P^2}{(P+\sigma^2)\sigma_z^2}}$ ,  $\alpha = \frac{P}{P+\sigma^2}$ ,  $S \sim \mathcal{N}(0, Q)$  and  $Z \sim \mathcal{N}(0, \sigma_z^2)$ . From now on, we drop the  $\epsilon$ 's in  $\kappa_w$  to improve clarity. Note that  $\alpha$  depends only on the assumed noise variance  $\sigma^2$  and not on  $\sigma_a^2$ .

From the observations  $[V', U, Y]$ , an optimal linear MMSE estimate of  $V$  is obtained. Similar to the definition in section III-B let  $\mathbf{\Lambda}$  be the covariance of  $[V', U, Y]^T$  and  $\mathbf{\Gamma}$  be the correlation between  $V$  and  $[V', U, Y]^T$ .

Hence

$$\mathbf{\Lambda} = \begin{pmatrix} \sigma_v^2 - \sigma_z^2 & \kappa(\sigma_v^2 - \sigma_z^2) & 0 \\ \kappa(\sigma_v^2 - \sigma_z^2) & P + \alpha^2 Q + \kappa_w^2 \sigma_v^2 & P + \alpha Q \\ 0 & P + \alpha Q & P + Q + \sigma_a^2 \end{pmatrix} \text{ and } \mathbf{\Gamma} = \begin{pmatrix} \sigma_v^2 - \sigma_z^2 & \kappa \sigma_v^2 & 0 \end{pmatrix}^T.$$

Then the distortion (in the presence of mismatch) is given by

$$D_a = \sigma_v^2 - \mathbf{\Gamma}^T \mathbf{\Lambda}^{-1} \mathbf{\Gamma} \quad (41)$$

This on further simplification yields

$$D_a = \left[ (Q\sigma^4 + (P(P+Q) + 2P\sigma^2 + \sigma^4)\sigma_a^2)\sigma_z^2 \right] \times \left[ P^2(P+Q) + P(P+Q)\sigma^2 + Q\sigma^4 + (P(2P+Q) + 3P\sigma^2 + \sigma^4)\sigma_a^2 \right]^{-1}. \quad (42)$$

Let us now look at a few special cases

### A. Hybrid Digital Analog Costa Coding

In this setup there is side information only at the transmitter. The distortion achievable for the user under SNR mismatch with the actual SNR greater than the designed SNR is obtained by setting  $\sigma_v = \sigma_z$  (42) and is given below.

$$D_{va} = [(Q\sigma^4 + (P(P+Q) + 2P\sigma^2 + \sigma^4)\sigma_a^2)\sigma_v^2] \times [P^2(P+Q) + P(P+Q)\sigma^2 + Q\sigma^4 + (P(2P+Q) + 3P\sigma^2 + \sigma^4)\sigma_a^2]^{-1} \quad (43)$$

The distortion in the source  $\mathbf{v}$  is shown in Fig.5 for a designed SNR of 10 dB as the actual channel SNR ( $10 \log 1/\sigma_a^2$ ) varies when the source and interference both have unit variance. It can be seen that the distortion in the source is smaller with the HDA Costa scheme than with the digital Costa scheme.

In some case, the distortion in estimating the interference at the receiver may also be of interest and can be obtained by estimating  $S$  from (38) and (40). The distortion is given below,

$$D_{sa} = [Q(P + \sigma^2)(P^2 + (2P + \sigma^2)\sigma_a^2)] \times [P^2(P+Q) + P(P+Q)\sigma^2 + Q\sigma^4(P(2P+Q) + 3P\sigma^2 + \sigma^4)\sigma_a^2]^{-1} \quad (44)$$

It can be seen from Fig. 5 that the distortion in estimating the interference is better for the digital scheme than for the HDA Costa scheme.

In [14], Sutivong *et al.* have studied a somewhat related problem. They consider the transmission of a digital source in the presence of an interference known at the transmitter with a fixed channel SNR. They study the optimal tradeoff between the achievable rate and the error in estimating the interference at the designed SNR. The main result is that we can get a better estimate of the interference if we transmit the digital source at a rate lesser than the channel capacity. There are important differences our work and that in [14]. First of all, we consider transmission of an analog source instead of a digital source. Secondly, we consider mismatch in the channel, i.e., our schemes are designed to be optimal at the designed SNR and as we move away from the designed SNR, we study the tradeoff between the error in estimating the interference and the distortion in the reconstruction of the analog source. This tradeoff is discussed below.

### B. Generalized HDA Costa Coding under channel mismatch

Next we analyze the performance of the generalized HDA Costa coding under channel mismatch. This case leads to some interesting analysis. By changing the source coding rate of the digital part  $R$ , we can tradeoff the distortion between the source and the interference in the presence of mismatch.

The different random variables and their relations are given below.

$$U = X + \alpha S + \kappa_1 E \quad (45)$$

$$Y = X + S + W_a \quad (46)$$

$$V = V^* + E \quad (47)$$

In the above equation  $\kappa_1 = \sqrt{\frac{P}{P+\sigma^2} \frac{(P+\sigma^2)-\sigma^2 2^{2R}}{\sigma_v^2}}$  (Again, we have dropped the  $\epsilon$  in the expression for  $\kappa_1$ .) From the above equations an estimate of  $S$  as well as  $V$  is obtained by taking a linear MMSE estimate as all the random variables are Gaussian. The resulting expressions of estimation error  $D_{sa}(R)$  and  $D_{va}(R)$  are given by

$$D_{va}(R) = [(\sigma_a^2(\sigma^2 + P)^2 + (\sigma^4 + \sigma_a^2 P)Q)\sigma_v^2] \times [(\sigma^2 + P)^2(\sigma_a^2 + P + Q) - 2^{2R}(\sigma^2 - \sigma_a^2)P(\sigma^2 + P + Q)]^{-1} \quad (48)$$

$$D_{sa}(R) = [(\sigma^2 + P)(2^{2R}(\sigma^2 - \sigma_a^2)P - (\sigma^2 + P)(\sigma_a^2 + P))Q] \times [2^{2R}(\sigma^2 - \sigma_a^2)P(\sigma^2 + P + Q) - (\sigma^2 + P)^2(\sigma_a^2 + P + Q)]^{-1} \quad (49)$$

The performance of the generalized HDA Costa scheme and HDA Costa scheme in relation to digital scheme is shown in fig. 5. For example in separation using digital Costa there is no improvement in our estimate of the analog source, but we get a better estimate of the interference as shown in fig.5. On the contrary for the HDA Costa scheme there is only a small improvement in the estimate of the interference but a good improvement in the estimate of the analog source. The generalized HDA also shows a difference in the estimate for the source and the interference for different rates  $R$  and performs as a digital Costa for the choice of  $R = C$  and as HDA Costa for the choice of  $R = 0$ . In effect we can tradeoff the estimation error in interference with the source by choosing different values of  $R$  when there is a channel mismatch.

### C. Hybrid Digital Analog Wyner Ziv

In this case the distortion could be obtained by setting  $Q = 0$  in (42). The actual distortion is given by

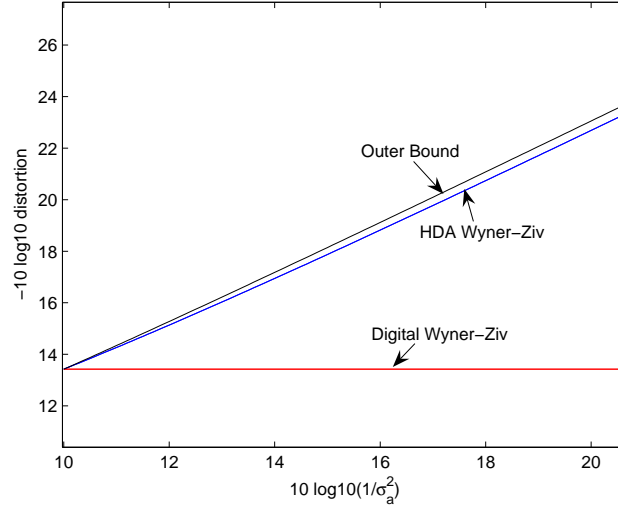


Fig. 6. Performance of the different Wyner-Ziv schemes for the joint source channel coding problem.

$$D_a = \frac{(P + \sigma^2)\sigma_a^2\sigma_z^2}{P^2 + (2P + \sigma^2)\sigma_a^2} \quad (50)$$

This is clearly better than  $\frac{\sigma_z^2\sigma^2}{P+\sigma^2}$  which is what is achievable with a separation based approach. However, we don't know if this is the optimal distortion that is achievable in the presence of channel mismatch. A simple lower bound on the achievable distortion in the presence of mismatch is to assume that the transmitter knows the channel SNR. Based on this we can analyze the gap in dB between the distortion of HDA Wyner Ziv scheme and the lower bound as follows.

The lower bound on  $D$  is given by

$$D_{lower} = \frac{\sigma_z^2}{1 + P/\sigma_a^2} \quad (51)$$

Now the gap between the analog Wyner-Ziv and the bound at high SNR can be easily calculated as  $\lim_{\sigma_a \rightarrow 0} \frac{D_{lower}}{D_a}$ . The gap in db,  $G$  is hence given by

$$G = 10 \log \left( \frac{P}{P + \sigma^2} \right) \quad (52)$$

For example, if our designed SNR is say 10 db, for high SNRs, we loose at most  $G = 0.41$  db which is fairly close to the outer bound as shown in Fig. 6.

## VI. DISTORTION EXPONENT FOR HDA COSTA AND WYNER-ZIV SCHEMES

In this section, we consider the performance of the HDA joint source-channel coding schemes for transmitting a Gaussian source through a Gaussian channel when the actual channel noise variance  $\sigma_a^2$  is not known, but it is known that the variance is always smaller than  $\sigma^2$ . Since we are interested in the performance of a single encoding scheme over a wide range of noise variances, a useful measure of performance is the rate of decay of the distortion as a function of the actual noise variance in the limit  $\sigma_a^2 \rightarrow 0$ . More precisely, we define a distortion exponent as

$$\zeta = \lim_{\sigma_a^2 \rightarrow 0} \frac{\log(D(\sigma_a^2))}{\log(\sigma_a^2)}$$

where  $D(\sigma_a^2)$  is the distortion when the noise variance is  $\sigma_a^2$ . Notice that this exponent is quite different from the distortion signal-to-noise ratio (SNR) exponent considered in [17], [18] for the case of slow fading channels. In [17] and [18], the rate of decay of distortion with average SNR is studied by allowing for a family of coding schemes, one for each average SNR. In contrast, we fix the encoding scheme here and consider the rate of decay with the actual channel SNR (i.e., there is no fading).

An upper bound on the achievable  $\zeta$  can be obtained by assuming that a genie informs the transmitter of  $\sigma_a$  and the transmitter chooses an optimal encoding scheme for a noise variance of  $\sigma_a^2$ . Let us assume a general case where there is an interference  $s$  ( $S \sim \mathcal{N}(0, Q)$ ) which is known at the transmitter and some side information  $\mathbf{v}'$  which is related to  $\mathbf{v}$  according to  $V = V' + Z$ , where  $Z \sim \mathcal{N}(0, \sigma_z^2)$  is known at the receiver. Then, the distortion for the genie-aided scheme is  $\frac{\sigma_z^2}{1 + \frac{P}{\sigma_a^2}}$  since an optimal Wyner-Ziv encoder followed by an optimal Costa encoder can be chosen. In this case, the distortion exponent is 1. Notice that in the absence of any side information the distortion for the genie-aided scheme is  $\frac{\sigma_z^2}{1 + \frac{P}{\sigma_a^2}}$  which also results in an exponent of 1. In the absence of any interference also, the achievable distortion is  $\frac{\sigma_z^2}{1 + \frac{P}{\sigma_a^2}}$  and the exponent is 1. Thus, for any single encoding scheme,  $\zeta \leq 1$  both in the presence and absence of interference and/or side information.

We will now consider the performance of the HDA schemes considered in Section IV. If a joint source channel coding scheme is designed to be optimal when the noise variance is  $\sigma^2$ , then the distortion when the noise variance is  $\sigma_a^2$  is given by (from (42) in Section V)

$$D(\sigma_a^2) = [(Q\sigma^4 + (P(P+Q) + 2P\sigma^2 + \sigma^4)\sigma_a^2)\sigma_z^2] \times [P^2(P+Q) + P(P+Q)\sigma^2 + Q\sigma^4 + (P(2P+Q) + 3P\sigma^2 + \sigma^4)\sigma_a^2]^{-1}. \quad (53)$$

We now consider two cases.

### A. Absence of Interference

When there is no interference at the transmitter,  $Q = 0$  and, hence, from (53), we can see that the optimal distortion can be obtained for a noise variance of  $\sigma^2$  and the optimal distortion exponent of  $\zeta = 1$  can be obtained. Thus, this scheme performs as well as the genie-aided receiver in the distortion exponent sense.

### B. Presence of Interference

In the presence of an interference,  $Q \neq 0$ , and from (53),  $\zeta$  can be seen to be zero. That is, some amount of residual interference is always present and, hence, in the high SNR limit ( $\sigma_a^2 \rightarrow 0$ ), the performance is dominated by this residual interference. However, if optimal performance is not desired when the noise variance is  $\sigma^2$ , then the optimal exponent of  $\zeta = 1$  can be obtained using a minor modification to the scheme discussed in Section IV. In the modified scheme, the auxiliary random variable  $U$  is generated as follows

$$U = X + S + \kappa_e V \quad (54)$$

Note that  $\alpha$  is chosen to be 1, which is clearly not optimal for a noise variance of  $\sigma^2$ . The side information  $V'$  is

$$V = V' + Z \quad (55)$$

and the received signal is

$$Y = X + S + W \quad (56)$$



Using arguments similar to those in Section IV,  $\kappa_e$  is chosen so as to satisfy  $I(U; Y, V') > I(U; S, V)$ . The required condition on  $\kappa_e$  can be obtained as follows

$$\begin{aligned} I(U; Y, V') &> I(U; S, V) \\ \Rightarrow h(U) - h(U|Y, V') &> h(U) - h(U|S, V) \\ \Rightarrow h(U|S, V) &> h(U|Y, V') \end{aligned} \quad (57)$$

Note that  $h(U|S, V) = h(X)$  and  $h(U|Y, V') = h(U - \eta Y - \kappa_e V'|Y, V')$ , where  $\eta = \frac{E[UY]}{E[Y^2]}$ . For this choice of  $\eta$ ,  $(U - \eta Y - \kappa_e V') \perp Y, V'$  and, hence,  $h(U|Y, V') = h(U - \eta Y - \kappa_e V')$ . Hence, we get the relation,

$$\begin{aligned} h(X) &> h(U - \eta Y - \kappa_e V') \\ \Rightarrow P &> E[(U - \eta Y - \kappa_e V')^2] \\ \Rightarrow \sqrt{\frac{P^2 + PQ - Q\sigma_z^2}{(P + Q + \sigma_z^2)\sigma_z^2}} &> \kappa_e \end{aligned}$$

Hence,  $\kappa_e$  can be chosen to be arbitrarily close to  $\sqrt{\frac{P^2 + PQ - Q\sigma_z^2}{(P + Q + \sigma_z^2)\sigma_z^2}}$ . Now  $\mathbf{x}$  is transmitted and  $\mathbf{y}$  is received. The optimal distortion is obtained as an MMSE estimate of  $\mathbf{v}$  from  $[\mathbf{y}, \mathbf{u}, \mathbf{v}']$ . The final distortion is given by

$$D(\sigma_a^2) = \frac{(P + Q)\sigma_a^2\sigma_z^2}{(P + Q)\sigma_a^2 + \kappa_e^2(P + Q + \sigma_a^2)\sigma_z^2} \quad (58)$$

It can be seen that as  $\sigma_a^2 \rightarrow 0$ ,  $D(\sigma_a^2) \propto \sigma_a^2$  and, hence,  $\zeta = 1$ .

## VII. APPLICATIONS TO TRANSMITTING A GAUSSIAN SOURCE WITH BANDWIDTH COMPRESSION

We now consider the problem of transmitting  $K$  samples of the i.i.d Gaussian source to a single user in  $N = \lambda K$  ( $\lambda < 1$ ) uses of an AWGN channel with noise variance  $\sigma^2$  where the transmit power is constrained to 1. There is no interference in the channel, but since  $\lambda < 1$ , we will see that the techniques described in the previous sections are useful for this problem.

There are at least three ways to achieve the optimal distortion in this case. One is to use a conventional separation based approach. The second one is to use superposition coding and the third one is to use Costa coding. Although, they are all optimal for the single user case, they perform differently when there is a mismatch in the channel SNR and, hence, the last two approaches are briefly described here.

*a) Superposition Coding:* Here we split the source in two parts and take  $N$  samples of the source  $\mathbf{v}$ , namely  $v_1^N$  and scale it by  $\sqrt{a}$  creating the systematic signal  $\mathbf{x}_1 = \sqrt{a}\mathbf{v}_1^N$ . We take the other  $K - N$  source samples  $v_{N+1}^K$  and use a conventional source encoder followed by a capacity achieving channel code resulting in the  $N$  dimensional vector  $\mathbf{x}_c = \mathcal{C}(\mathcal{Q}(v_{N+1}^K))$ , where  $\mathcal{C}$  denotes a channel encoding operation and  $\mathcal{Q}$  denotes a source encoding operation. Then  $\mathbf{x}_c$  is normalized so that the average power is  $\sqrt{1 - a}$ . The overall transmitted signal is  $\mathbf{x} = \mathbf{x}_s + \mathbf{x}_c$  and the received signal is  $\mathbf{y} = \mathbf{x} + \mathbf{w}$ . At the receiver, the digital part is first decoded assuming the systematic (analog) part is noise and then  $\mathbf{x}_c$  is subtracted from  $\mathbf{y}$ . Then an MMSE estimate of  $v_1^N$  is formed. For the optimal choice of  $a$ , the optimal overall distortion can be obtained given by

$$a_{sup}^* = \sigma^2 \left[ \left( 1 + \frac{1}{\sigma^2} \right)^\lambda - 1 \right] \text{ and } D_{sup}^* = \frac{1}{\left( 1 + \frac{1}{\sigma^2} \right)^\lambda} \quad (59)$$

which is the optimal distortion.

*b) Digital Costa Coding:* We split the source exactly as in the previous case and one stream is formed as  $\mathbf{x}_s = \sqrt{a}\mathbf{v}_1^N$ . However, here the digital part assumes that  $\mathbf{x}_s$  is interference and uses Costa coding to produce  $\mathbf{x}_c$  with power  $1 - a$  as shown in Fig. 7. In Costa coding, we define an auxiliary random variable  $\mathbf{u} = \mathbf{x}_c + \alpha_1 \mathbf{x}_s$  where  $\alpha_1 = \frac{1-a}{1-a+\sigma^2}$  is the optimum scaling coefficient. At the receiver, the digital part is decoded which means that  $\mathbf{u}$  can be obtained. In spite of knowing  $\mathbf{u}$  exactly, the optimal estimate of  $v_1^N$  is obtained by simply treating  $\mathbf{x}_c$  as noise since for the optimal choice of  $\alpha_1$ ,  $\mathbf{x}_c = \mathbf{u} - \alpha_1 \mathbf{x}_s$  and  $v_1^N$  are uncorrelated. Therefore, an MMSE estimate of  $v_1^N$  is formed assuming  $\mathbf{x}_c$  were noise. Hence, the overall distortion becomes

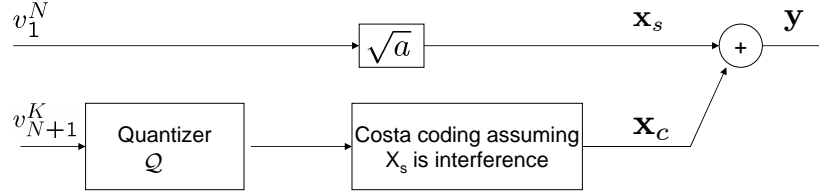


Fig. 7. Encoder model using Costa coding for single user

$$D = \frac{\lambda}{1 + \frac{a}{1-a+\sigma^2}} + \frac{1-\lambda}{\left(1 + \frac{1-a}{\sigma^2}\right)^{\lambda/(1-\lambda)}} \quad (60)$$

Again, minimizing  $D$  w.r.t.  $a$  gives

$$a_{costa}^* = (1 + \sigma^2) \left[ 1 - \frac{1}{\left(1 + \frac{1}{\sigma^2}\right)^\lambda} \right] \text{ and } D_{costa}^* = \frac{1}{\left(1 + \frac{1}{\sigma^2}\right)^\lambda} \quad (61)$$

which is the best possible distortion.

c) *Hybrid Digital Analog Costa Coding*: For the case of  $\lambda = 0.5$ , the digital Costa coding part can be replaced by a hybrid digital analog (HDA) Costa coding. We refer to such a scheme as HDA Costa coding. The same power allocation however, remains the same and hence, we can simply use  $a_{Costa}^*$  without the need to differentiate the digital and HDA Costa coding. It is quite straightforward to show that  $a_{Costa}^* > a_{sup}^*$  for  $\lambda < 1$ . Hence, the Costa coding approach allocates higher power to the systematic part than the superposition approach, since the systematic part is treated as interference.

#### A. Performance in the presence of SNR mismatch

Now, we consider the same set up as above, but when the actual channel noise variance is  $\sigma_a^2$ , whereas the designed noise variance is  $\sigma^2$ .

Case 1:  $\sigma_a^2 > \sigma^2$

The distortion for the superposition code can be computed to be the sum of the distortions in the systematic part and the digital part. When  $\sigma_a^2 > \sigma^2$ , the digital part cannot be decoded and, hence, we assume that the distortion in the digital part is the variance of the source, 1.

$$D_{sup} = \frac{\lambda}{1 + \frac{a_{sup}^*}{1-a_{sup}^*+\sigma_a^2}} + (1-\lambda) \cdot 1 \quad (62)$$

Both the digital and HDA Costa coding schemes perform identically when  $\sigma_a^2 > \sigma^2$  and the distortion for the Costa code can be computed to be

$$D_{digCosta} = D_{HDACosta} = \frac{\lambda}{1 + \frac{a_{Costa}^*}{1-a_{Costa}^*+\sigma_a^2}} + (1-\lambda) \cdot 1 \quad (63)$$

Case 2:  $\sigma_a^2 < \sigma^2$  In this case, the digital part can be decoded exactly and, hence, the distortion for superposition coding is

$$D_{sup} = \lambda \frac{1}{1 + \frac{a_{sup}^*}{\sigma_a^2}} + (1-\lambda) \frac{1}{\left(1 + \frac{1-a_{sup}^*}{a_{sup}^*+\sigma^2}\right)^{\lambda/(1-\lambda)}} \quad (64)$$

For digital Costa coding, the decoder first decodes the digital part when the auxiliary random variable  $\mathbf{u}$  is perfectly known. In the case when  $\sigma_a^2 \neq \sigma^2$ , the receiver must form the MMSE estimate of  $v_1^N$  from the channel observation

$\mathbf{y}$  and  $\mathbf{u}$ . Therefore, the overall distortion is

$$D_{\text{digCosta}} = \lambda \left( 1 - [\sqrt{a_{\text{Costa}}^*} \alpha \sqrt{a_{\text{Costa}}^*}] \times \begin{bmatrix} 1 + \sigma_a^2 & 1 - a_{\text{Costa}}^* + \alpha a_{\text{Costa}}^* \\ 1 - a_{\text{Costa}}^* + \alpha a_{\text{Costa}}^* & 1 - a_{\text{Costa}}^* + \alpha^2 a_{\text{Costa}}^* \end{bmatrix}^{-1} \times \begin{bmatrix} \sqrt{a_{\text{Costa}}^*} \\ \alpha \sqrt{a_{\text{Costa}}^*} \end{bmatrix} \right) + (1 - \lambda) \frac{1}{\left( 1 + \frac{1 - a_{\text{Costa}}^*}{a_{\text{Costa}}^* + \sigma_a^2} \right)^{\lambda/(1-\lambda)}} \quad (65)$$

For the HDA Costa coding, we can decode  $\mathbf{u}$  and form MMSE estimates of  $v_1^N$  and  $v_{N+1}^K$  separately and, hence, the overall distortion is given by

$$D_{\text{HDACosta}} = \lambda \left( 1 - [\sqrt{a_{\text{Costa}}^*} \alpha \sqrt{a_{\text{Costa}}^*}] \times \begin{bmatrix} 1 + \sigma_a^2 & 1 - a_{\text{Costa}}^* + \alpha a_{\text{Costa}}^* \\ 1 - a_{\text{Costa}}^* + \alpha a_{\text{Costa}}^* & 1 - a_{\text{Costa}}^* + \alpha^2 a_{\text{Costa}}^* + \kappa^2 \end{bmatrix}^{-1} \times \begin{bmatrix} \sqrt{a_{\text{Costa}}^*} \\ \alpha \sqrt{a_{\text{Costa}}^*} \end{bmatrix} \right) + (1 - \lambda) (1 - [0 \ \kappa] \times \begin{bmatrix} 1 + \sigma_a^2 & 1 - a_{\text{Costa}}^* + \alpha a_{\text{Costa}}^* \\ 1 - a_{\text{Costa}}^* + \alpha a_{\text{Costa}}^* & 1 - a_{\text{Costa}}^* + \alpha^2 a_{\text{Costa}}^* + \kappa^2 \end{bmatrix}^{-1} \times \begin{bmatrix} 0 \\ \kappa \end{bmatrix}) \quad (66)$$

The performance of the superposition scheme, digital Costa and HDA Costa scheme are shown for an example with  $\lambda = 0.5$  in Fig. 8. The designed SNR is defined as  $10 \log_{10} \frac{1}{\sigma_a^2}$  whereas the actual SNR is defined as  $10 \log_{10} \frac{1}{\sigma_a^2}$ . In the example, the designed SNR is fixed at 10dB and the actual SNR is varied from 0 dB to 20 dB. It can be seen that the Costa coding approach is better than superposition coding when  $\sigma_a^2 > \sigma^2$  and worse for the other case. The HDA Costa coding scheme performs the best over the entire range of SNRs.

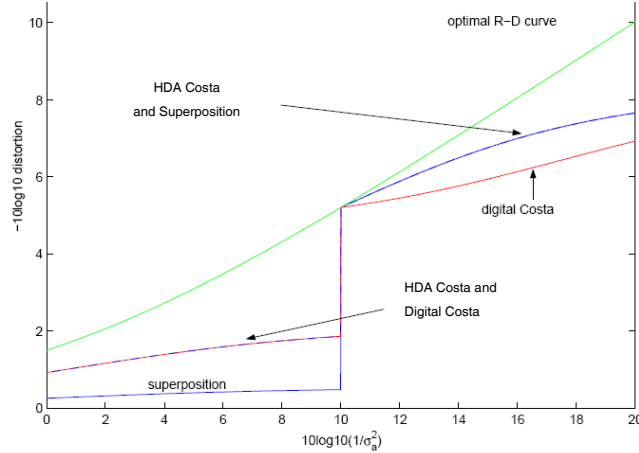


Fig. 8. Performance of different schemes for the source splitting approach for the bandwidth compression problem with SNR mismatch.

## VIII. APPLICATIONS TO BROADCASTING WITH BANDWIDTH COMPRESSION

We now consider the problem of transmitting  $K = 2N$  samples of a unit variance Gaussian source  $\mathbf{v}$  in  $N$  uses of the channel to two users through AWGN channels with noise variances  $\sigma_1^2$  (weak user) and  $\sigma_2^2$  (strong user) with  $\sigma_1 > \sigma_2$ . The channel has the power constraint  $P = 1$ . We are interested in joint source channel coding schemes

that provide a good region of pairs of distortion that are simultaneously achievable at the two users. This problem was considered in [1, 5, 8]. The best known region to date is given by the schemes therein.

Notice that when we design a source channel coding scheme to be optimal for the weak user, the strong user operates under the situation of SNR mismatch explained in Section VII-A with  $\sigma_2^2 = \sigma_a^2 < \sigma^2 = \sigma_1^2$ . Similarly, when the system is designed to be optimal for the strong user, for the weak user  $\sigma_1^2 = \sigma_a^2 > \sigma^2 = \sigma_2^2$ . Motivated by the fact that for  $\lambda = 0.5$ , the HDA Costa coding scheme performs the best, we propose a scheme which is shown in Fig. 9.

There are three layers in the proposed coding scheme. The first layer is the systematic part where  $N$  out of the  $K$  samples of the source are scaled by  $\sqrt{a}$ . Let us call this as  $\mathbf{x}_s = \sqrt{a}v_1^N$ . The other  $K - N$  samples of the Gaussian source are hybrid digital analog Costa coding, treating  $\mathbf{x}_s$  as the interference and transmits the signal  $\mathbf{x}_1$  with power  $b$  in the second layer. So  $\mathbf{x}_1 = \mathbf{u}_1 - \alpha_1 \mathbf{x}_s - \kappa_c v_{N+1}^K$ , where  $\alpha_1$  and  $\kappa_c$  are the optimal scaling coefficient to be used in the hybrid digital analog Costa coding process and  $\mathbf{u}_1$  is the auxiliary variable. This layer is meant to be decoded by the weak user and, hence, the scaling factor  $\alpha_1$  is set to be  $b/(b + c + \sigma_1^2)$ . That is, this layer sees the third layer also as *independent* noise.

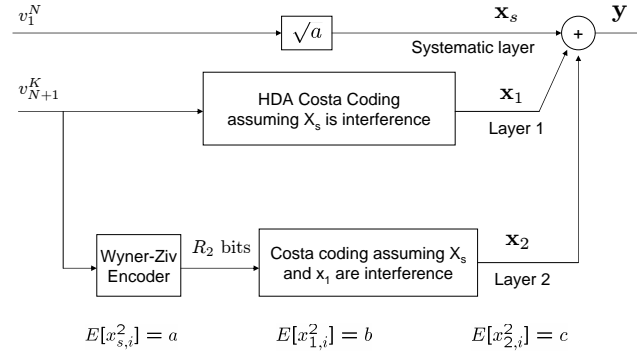


Fig. 9. Encoder model using Costa coding

The third layer is first Wyner Ziv coded at a rate  $R_2$  assuming the estimate of  $v_{N+1}^K$  at the receiver as side information. The Wyner-Ziv index is then encoded using digital Costa coding assuming  $\mathbf{x}_s$  and  $\mathbf{x}_1$  are interference and uses power  $c = 1 - a - b$ . Therefore,  $\mathbf{x}_2 = \mathbf{u}_2 - \alpha_2(\mathbf{x}_s + \mathbf{x}_1)$ . This layer is meant for the strong user and, hence, the scaling factor  $\alpha_2 = c/(c + \sigma_2^2)$ . We then transmit  $\mathbf{x} = \mathbf{x}_s + \mathbf{x}_1 + \mathbf{x}_2$ .

At the receiver, from the second layer an estimate of  $v_{N+1}^K$  is obtained. This estimate acts as side information that can be used in refining the estimate of  $v_{N+1}^K$  for the strong user using the decoded Wyner-Ziv bits. The Wyner-Ziv bits are decoded from the third layer by Costa decoding procedure.

The users estimate the systematic part  $v_1^N$  and non-systematic part  $v_{N+1}^K$  by MMSE estimation from the received  $\mathbf{y}$ , the decoded  $\mathbf{u}_1$  and  $\mathbf{u}_2$ . So the overall distortion seen at the weak user is

$$D_1 = \frac{1}{2} \frac{1}{1 + \frac{a}{b+c+\sigma_1^2}} + \frac{1}{2} \frac{1}{1 + \frac{b}{c+\sigma_1^2}}$$

The distortion for the strong user is given by

$$D_2 = \frac{1}{2} \left( 1 - [\sqrt{a} \ \alpha_1 \sqrt{a}] \begin{bmatrix} 1 + \sigma_2^2 & b + \alpha_1 a \\ b + \alpha_1 a & b + \alpha_1^2 a + \kappa^2 \end{bmatrix}^{-1} \times \right. \\ \left. \begin{bmatrix} \sqrt{a} \\ \alpha_1 \sqrt{a} \end{bmatrix} \right) + \frac{1/2}{1 + \frac{c}{\sigma_2^2}} (1 - [0 \ \kappa] \times \\ \left. \begin{bmatrix} 1 + \sigma_2^2 & b + \alpha_1 a \\ b + \alpha_1 a & b + \alpha_1^2 a + \kappa^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \kappa \end{bmatrix} \right) \quad (67)$$

The corner points of the distortion region corresponding to being optimal for the strong and weak user respectively, can be obtained by setting  $c = 0$  and  $b = 0$ , respectively.

The distortion region for this scheme for the case of  $\sigma_1^2 = 0$  dB and  $\sigma_2^2 = 5$  dB is shown in Fig. 10. The distortion region for three other schemes are also shown. They are the scheme proposed by Mittal and Phamdo in [1], a different broadcasting scheme which uses digital Costa coding in both the layers proposed in [8] (details can be found there) and the broadcast scheme with one layer of superposition coding and one layer of digital Costa coding considered in [5, 8]. This scheme currently appears to be the best known scheme. Notice that in the third layer, instead of using a separate Wyner-Ziv encoder followed by a Costa code, we could have used the HDA scheme discussed in Section IV with identical results.

The proposed broadcast scheme in Fig. 9 significantly outperforms the scheme in Mittal and Phamdo and the digital Costa based broadcast scheme for this example. The corner points of this scheme also coincide with those of the best known schemes reported in [5, 19].

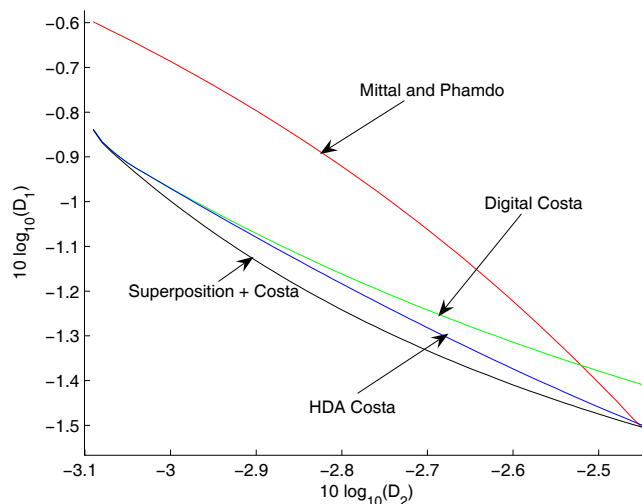


Fig. 10. Distortion regions of the different schemes for broadcasting with bandwidth compression.

## IX. CONCLUSION AND FUTURE WORK

We discussed hybrid digital analog version of Costa coding and Wyner-Ziv coding for transmitting an analog Gaussian source through an AWGN channel in the presence of an interferer known only to the transmitter and side information available only to the receiver respectively. These schemes are closely related to the schemes by Reznicek and Zamir [2] and [7], but make the auxiliary random variable model more explicit. We also showed that there are infinitely many schemes that are optimal for this problem, extending the work of Bross, Lapidot and Tinguely [6] to the side information case. The HDA coding schemes have advantages over strictly digital schemes when there is a mismatch in the channel SNR. This makes them also useful for broadcasting a Gaussian source to two users with different SNRs.

## REFERENCES

- [1] U. Mittal and N. Phamdo, "Hybrid digital-analog HDA joint source-channel codes for broadcasting and robust communications," *IEEE Tran. Info. Theory*, vol. 48, no. 5, pp. 1082–1102, May 2002.
- [2] Z. Reznicek, M. Feder, and R. Zamir, "Distortion bounds for broadcasting with bandwidth expansion," *IEEE Tran. Info. Theory*, vol. 52, no. 8, pp. 3778–3788, August 2006.
- [3] T. J. Goblick Jr., "Theoretical limitations on the transmission of data from analog sources," *IEEE Tran. Info. Theory*, vol. 11, no. 10, pp. 558–567, Oct 1965.
- [4] M. Gastpar, B. Rimoldi and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Tran. Info. Theory*, vol. 49, no. 5 pp. 1147–1158, May 2003.
- [5] V. M. Prabhakaran, R. Puri, and K. Ramachandran, "Hybrid Analog-Digital Strategies for Source-Channel Broadcast" *43rd Allerton Conference on Communication, Control and Computing*, Allerton, IL, September 2005.

- [6] S. Bross, A. Lapidoth, and S. Tinguely, "Superimposed Coded and Uncoded Transmissions of a Gaussian Source over the Gaussian Channel," *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Seattle, USA, 2006.
- [7] Y. Kochman and R. Zamir, "Analog Matching of Colored Sources to Colored Channels," *ISIT 2006*, Seattle, USA, July 2006.
- [8] K. R. Narayanan, M. P. Wilson and G. Caire, "Hybrid Digital and Analog Costa Coding for Broadcasting with Bandwidth Compression," Wireless Communications Lab Technical report, TR-06-107, Texas A&M University, <http://wcl3.tamu.edu/research.html>, August 25, 2006
- [9] M. Costa, "Writing on Dirty Paper," *IEEE Tran. Info. Theory*, vol. 29, No. 3, pp. 439-441, May 1983.
- [10] S. Shamai, S. Verdu, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Info. Theory*, vol. 44, pp. 564-579, March 1998.
- [11] N. Merhav, and S. Shamai, "On joint source-channel coding for the Wyner-Ziv source and the Gel'fand-Pinsker channel," *IEEE Trans. Info. Theory*, vol. 49, pp. 2844-2855, November 2003.
- [12] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Info. Theory*, vol. 48, pp. 1250-1276, June 2002.
- [13] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Info. Theory*, vol. 51, pp. 3401-3416, October 2005.
- [14] A. Sutivong, M. Chiang, T. M. Cover, and Y. Kim, "Channel Capacity and State Estimation for State-Dependent Gaussian Channels," *IEEE Tran. Info. Theory*, vol. 51, No. 4, pp. 1486-1495, April 2005.
- [15] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, 2006
- [16] A. D. Wyner and J. Ziv, "The rate distortion function for source coding with side information at the decoder," *IEEE Tran. Info. Theory*, vol. IT-22, no. 1, pp. 1-10, Jan. 1976.
- [17] J. N. Laneman, E. Martinian, G. W. Wornell, and J. G. Apostolopoulos, "Source channel diversity for parallel channels", *IEEE Transactions on Information Theory*, vol. IT-51, no. 10, pp. 3518-3539, Oct 2005.
- [18] T. Holliday and A. J. Goldsmith, "Joint source and channel coding for mimo systems", *Allerton Conf. Commun. Control and Computing*, pp. 1302-1311, Monticello IL, USA, Oct. 2004.
- [19] K. R. Narayanan, G. Caire and M. P. Wilson, "Duality Between Broadcasting with Bandwidth Expansion and Bandwidth Compression", *Proc. Intl. Symp. Info. Theory*, Nice, France 2007.